Genome screening of coeliac disease

S Popat, S Bevan, C P Braegger, A Busch, D O'Donoghue, K Falth-Magnusson,
A Godkin, L Hogberg, G Holmes, K B Hosie, P D Howdle, H Jenkins, D Jewell,
S Johnston, N P Kennedy, P Kumar, R F A Logan, A H G Love, M N Marsh, C J Mulder,
K Sjoberg, L Stenhammar, J Walker-Smith, R S Houlston

....

Coeliac disease is caused by T cell sensitisation of the intestine to cereal prolamins, which results in a range of mucosal abnormalities that may lead to malabsorption.[1] The population prevalence in western countries is ~1 in 200.[2–4] Evidence for an inherited predisposition to coeliac disease comes from studies of first degree relatives of patients and studies of twins.[5] [6] A strong association is seen between coeliac disease and the HLA-DQ (α1*05, β1*02) heterodimer (DQ2) which is present in approximately 95% of patients,[7–9] compared with 20-30% of healthy subjects.[10] [11] The difference in concordance rates between monozygotic twins and HLA identical sibs (80-100% *v* 25%) implicates non-HLA genes in the genetic predisposition to coeliac disease.[11] The overall relative risk in sibs is at least 20 and is therefore fourfold higher than that attributable to HLA alone under model of inheritance.[7] Genome linkage searches carried out on Irish,[12] Italian,[13] and UK[14] coeliac disease families have identified a number of potential sites for the location of non-HLA linked genes. The putative candidate loci detected in the three studies are, however, largely inconsistent and the findings have not been replicated in other populations.[15–19] Here we report the results of a genome screen of 24 multiplex families with coeliac disease and discuss the findings of this study in relation to previously published analyses.

METHODS AND RESULTS

Twenty-four families with two or more members affected with coeliac disease were recruited for this study (fig 1). Nine of these families (Nos 1, 3, 4, 6, 30, 32, 39, 43, and 44) have been used in a previous study of candidate regions.[16] All the families were of northern European ancestry. Twelve of the families were recruited from the UK, nine from Sweden, two from Switzerland, and one from The Netherlands. All affected family members had symptomatic disease and were diagnosed according to the revised criteria formulated by the European Society for Paediatric Gastroenterology, Hepatology, and
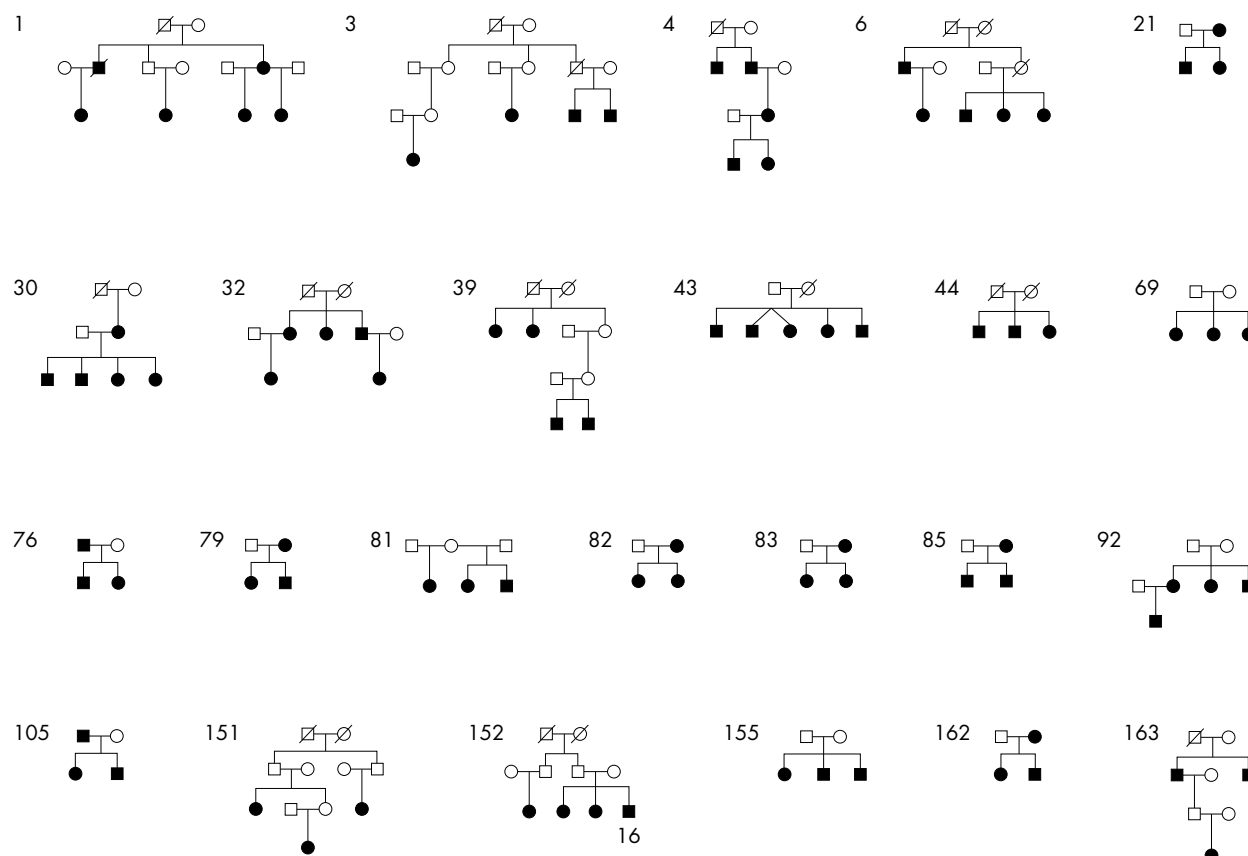

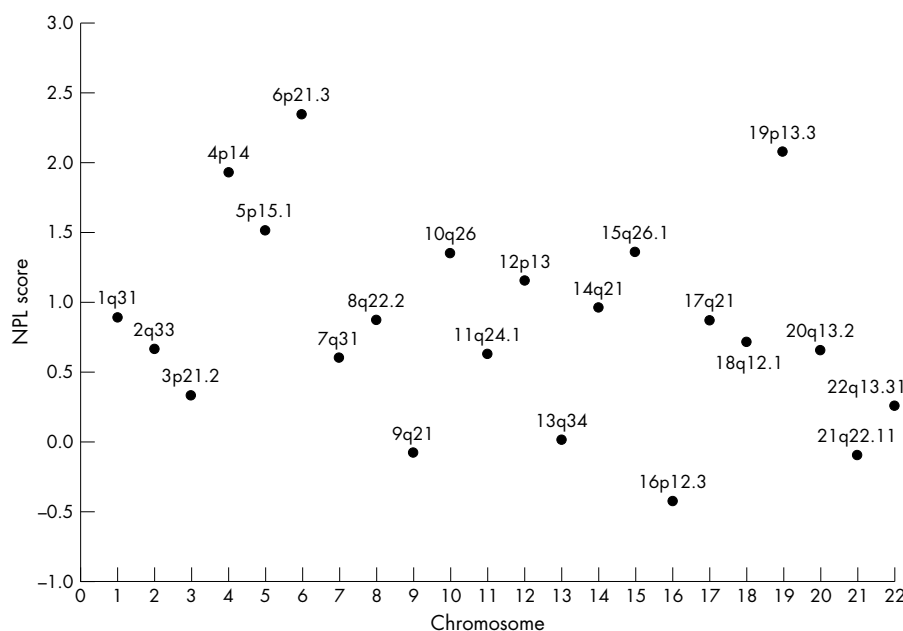**Figure 1** Pedigree structures of the families studied.

Figure 2 NPL scores by chromosome.

Nutrition (ESPGHAN).[20] Families consisted of 88 affected subjects, 36 male and 52 female, giving a male to female ratio of 1:1.5. The age at diagnosis ranged from 9 months to 74 years with a mean of 25 years. Blood samples and clinical data were all collected with informed consent and ethical review board approval. All family members were typed for the HLA-DQ A1*0501 and B1*0201 alleles by PCR-SSP[21] Ninety-nine percent of affected subjects in the families studied had HLA class II genotypes compatible with possession of the DQA1*05, B1*02 alleles.

Fluorescent genotyping[16] was performed at 217 microsatellite markers covering every autosome at an average distance of 16.4 cM. Genetic linkage was assessed by determining multipoint NPL (non-parametric linkage) scores using the GENEHUNTER-PLUS program.[22] Multipoint NPL scores are model independent and effectively measure the extent to which marker allele sharing by descent between affected subjects is greater than expected under random segregation. Fig 2 shows the distribution of NPL scores by chromosome. The strongest region of linkage was observed at chromosome 6p21.3, at the site of HLA (nominal p value 0.01). As expected there was evidence of excess sharing at both HLA and close flanking markers and this underlies the 6p linkage. There was no evidence of linkage to any region telomeric to HLA. Outside chromosome 6p21 the strongest linkage was obtained at chromosome 4p14 (maximal at D4S405, nominal p value 0.03) and at chromosome 19p13.3 (maximal between D19S424 and D19S894, nominal p value 0.02).

DISCUSSION

Three genome wide linkage searches of coeliac disease have been published,[12–14] but no regions of significant linkage have been reported consistently. The first of these was reported by Zhong et al.[12] Linkage of coeliac disease to five chromosome regions outside HLA was detected: 6p23 (telomeric to HLA), 7q31.3, 11p11, 15q26, and 22cen. In addition to these regions, there was also some evidence of linkage to chromosome 19p13.3 (p=0.05) as seen in our study. The second study was reported by Greco et al.[13] In addition to HLA, there was some evidence for linkage to 5qter and 11qter, but no evidence to support the notion that there is a locus on chromosome 6p distinct from HLA as proposed by Zhong et al.[12] A follow up study of candidate regions reported by Greco et al[19] failed to

confirm linkage to chromosome 11qter, but gave some support for linkage to chromosome 5q (maximum MLS 2.9). We were unable to find any evidence of linkage to chromosome 5q region in our dataset. The genome wide study recently reported by King et al[14] was only based on analysis of 16 UK multiplex coeliac disease families. Two regions, chromosomes 10q23.1 and 16q23.3, provided evidence of linkage in multipoint analyses using five markers (nominal p values of 0.006). There was limited evidence for linkage to HLA in the study (lod score <1.2).

In addition to a failure to confirm linkage to chromosome 6p23, neither our study nor that of Greco et al[13] and King et al[14] found evidence for linkage to chromosome 11p11, as originally proposed by Zhong et al.[12] The proposition of an additional chromosome 6p locus made by Zhong et al[12] was based on typing 30 of 45 affected sib pairs at D6S259, and flanking markers within 1 cM showed no significant evidence for linkage. In our study, there was no evidence for linkage to this region. Similarly, there is no support for linkage of coeliac disease to a locus on chromosome 6 telomeric to HLA in the studies reported by Greco et al[13] and Brett et al.[15]

Although we found no significant evidence for linkage to chromosomes 5qter or 11qter as initially suggested by Greco et al,[13] we cannot preclude these loci as sites of a non-HLA linked locus. We did, however, find some evidence for linkage at chromosome 19p13.3, raising the possibility that this region may define an additional locus.

While coeliac disease is oligogenic, the simple HLA association coupled with the fact that the environmental trigger is identical in all subjects theoretically makes genetic analysis of the trait more amenable to dissection than other complex disorders, such as non-insulin dependent diabetes mellitus or asthma. As expected, our study confirms HLA as a risk factor for coeliac disease. Two other regions showed evidence for comparable linkage, chromosomes 19p13.3 and 4p14. However, neither attains the level of statistical significance desired for genome wide linkage searches and hence require confirmation.

The genetic analysis of coeliac disease potentially has many advantages over other complex traits, both from the perspective of the magnitude of the genetic trait and the ease of defining affected status. However, the studies reported to date suggest that identifying the non-HLA linked component may not be straightforward. It is possible that the discrepancies in

findings between the four studies are because of differences in the contribution of genetic factors to coeliac disease in the different populations analysed. Alternatively and specifically related to the plethora of significant findings reported by Zhong et al,[12] this might reflect in part the structure of the families studied. Specifically, 31 of the ASPs belonged to three families. The inclusion of large sibships in which no typing information is available may well have led to artificially inflated support for specific regions.

If disease susceptibility is the result of more than one gene, the power to detect linkage depends on the relative contribution to the overall familial risk made by each locus and how the different loci interact. The power of this and previously published studies[12–14] to detect linkage and the effect of varying heterogeneity between zero and 75% was assessed by non-parametric means using the program ALLEGRO.[23] In the absence of a definitive model for the mode of inheritance of coeliac disease, genotypes were simulated under models chosen on the basis that the sib relative risk conferred by the disease gene was ~3.3. This follows from the assumptions that the overall disease prevalence is 0.005 and that the HLA linked and unlinked genes interact multiplicatively.[7] The models were a common recessive disease gene with frequency 0.2, a rare recessive disease gene with frequency 0.05, a common dominant disease gene with frequency 0.05, and a rare dominant disease gene with frequency 0.003. Even under the assumption of homogeneity, the power of each of the studies to show significant linkage over all models is marginal, especially in the study reported by Zhong et al,[12] which has a maximum power of 56% to detect linkage under any of the models. The study by King et al[14] is consistently less powered than the present one, and this reflects the small sample size used. In all studies, the power to detect linkage falls markedly in the presence of heterogeneity, with a power of less than 20% to detect linkage under 50% heterogeneity, under all models. This suggests that current genome wide linkage searches of coeliac disease are underpowered. If coeliac disease is heterogeneous, subcategorisation of disease by clinical phenotype or HLA haplotype might provide larger familial risks and hence increased power. Furthermore, the use of isolated groups such as the Finnish or Sardinian populations with limited genetic heterogeneity can be advantageous because of a founder effect. The optimal linkage design for the detection of a non-HLA predisposition gene for coeliac disease is not as simple as for a rare high penetrance gene. Aside from our study and that of King et al,[14] the two previous genome wide searches[12 13] have been based on affected sibs, excluding four of the 15 families in the study reported by Zhong et al.[12] Families comprising three or more cases are likely to be more powerful for linkage than affected sib pairs or other two case families. However, increasing the number of affected subjects in a family may not necessarily represent an efficient strategy, depending on the genetic model. On the basis of segregation studies, homozygosity at an HLA unlinked locus is a perquisite for developing coeliac disease.[24–26] Hence, there is a possibility that in any large family two copies of a mutation may be segregating if the population frequency of the disease allele is high. Practically, the composition of families used in linkage analysis will always depend on what can efficiently be collected. It is clear that further studies, ideally commensurate with the detection of linkage under heterogeneity, are required to support or refute the possible location of putative non-HLA linked genes.

ACKNOWLEDGEMENTS

. . . .
Authors' affiliations
S Popat, S Bevan, R S Houlston, Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK
C P Braegger, University Children's Hospital, Steinwiesstrass 75, Zurich, Switzerland
A Busch, Kinderpoliklinik, Eberhard-Karls-Universitat, Tubingen, Germany
D O'Donoghue, Gastroenterology and Liver Unit, St Vincent's Hospital, Dublin, Ireland
K Falth-Magnusson, Department of Paediatrics, Linkoping University, Sweden
A Godkin, D Jewell, Gastroenterology Unit, The Radcliffe Infirmary, Oxford, UK
L Hogberg, L Stenhammar, Department of Paediatrics, Norrkoping Hospital, Sweden
G Holmes, Department of Medicine, Derbyshire Royal Infirmary, Derby, UK
K B Hosie, University Surgical Unit, Northern General Hospital, Sheffield, UK
P D Howdle, Division of Medicine, St James's University Hospital, LeedU
S Johnston, A H G Love, Department of Gastroenterology, Belfast City Hospital Trust, Belfast, UK
N P Kennedy, Department of Clinical Medicine, Trinity Centre, St James' Hospital, Dublin, Ireland
P Kumar, Department of Gastroenterology, St Bartholomew's Hospital, London, UK
R F A Logan, Department of Public Health and Epidemiology, University Hospital, Nottingham, UK
M N Marsh, University Department of Medicine, Hope Hospital, Salford, Manchester, UK
C J Mulder, Hepatogastroenterology, Rijnstate Hospital, Arnhem, The Netherlands
K Sjoberg, Department of Medicine, University of Lund, Malmo, Sweden
J Walker-Smith, University Department of Paediatric Gastroenterology, Royal Free Hospital, London, UK

Correspondence to: Dr S Popat, Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK; s.popat@icr.ac.uk

REFERENCES

1 **Marsh MN**. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 1992;**102**:330-54.
2 **Csizmadia CG**, Mearin ML, von Blomberg BM, Brand R, Verloove-Vanhorick SP. An iceberg of childhood coeliac disease in the Netherlands. *Lancet* 1999;**353**:813-14.
3 **Feigherty C**. Coeliac disease. *BMJ* 1999;**319**:236-9.
4 **Hin H**, Bird G, Fisher P, Mahty N, Jewell D. Coeliac disease in primary care: case finding study. *BMJ* 1999;**318**:164-7.
5 **Ellis A**. Coeliac disease: previous family studies. In: McConnell RB, ed. *The genetics of coeliac disease*. Lancaster: MTP, 1981:197-9.
6 **Polanco I**, Biemond I, van Leeuwen A, Schreuder I, Meera Khan P, Guerrero J, D'Amaro J, Vasquez C, van Rood JJ, Pena AS. Gluten sensitive enteropathy in Spain: genetic and environmental factors. In: McConnell RB, ed. *The genetics of coeliac disease*. Lancaster: MTP, 1981:211-31.
7 **Houlston RS**, Ford D. Genetics of coeliac disease. *Q J Med* 1996;**89**:737-43.
8 **Kagnoff MF**, Harwood JI, Bugawan TL, Bugawan TL, Erlich HA. Structural analysis of the HLA-DR, -DQ, and -DP alleles on the celiac disease-associated HLA-DR3 (DRw17) haplotype. *Proc Natl Acad Sci USA* 1989;**86**:6274-8.
9 **Tosi R**, Vismara D, Tanigaki N, Ferrara GB, Cicimarra F, Buffolano W, Follo D, Auricchio S. Evidence that celiac disease is primarily associated with a DC locus allelic specificity. *Clin Immunol Immunopathol* 1983;**28**:395-404.
10 **Sollid LM**. Molecular basis of celiac disease. *Annu Rev Immunol* 2000;**18**:53-81.
11 **Sollid LM**, Thorsby E. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* 1993;**105**:910-22.
12 **Zhong F**, McCombs CC, Olson JM, Elston RC, Stevens FM, McCarthy CF, Michalski JP. An autosomal screen for the genes that predispose to celiac disease in the western counties of Ireland. *Nat Genet* 1996;**14**:329-33.
13 **Greco L**, Corazza G, Barbon MC, Clot F, Fulchignoni-Lataud MC, Percopo S, Zavattari P, Bouguerra F, Dib C, Tosi R, Trocone R, Ventura A, Mantavoni W, Magazzu G, Gatti R, Lazzari R, Giunta A, Perri F, Iacono G, Cardi E, de Virgilis S, Cataldo F, De Angelis G, Musumeci S, Ferrari R, Balli F, Bardella MT, Volta U, Catassi C, Torre G, Eliaou JF, Serre JL, Clerget-Darpoux F. Genome search in celiac disease. *Am J Hum Genet* 1998;**62**:669-75.
14 **King AL**, Yiannakou JY, Brett PM, Curtis D, Morris MA, Dearlove AM, Rhodes M, Rosen-Bronson S, Mathew C, Ellis HJ, Ciclitira PJ. A genome-wide family-based linkage search of coeliac disease. *Ann Hum Genet* 2000;**64**:497-0.
15 **Brett PM**, Yiannakou JY, Morris MA, Bronson SR, Mathew C, Curtis D, Ciclitira PJ. A pedigree-based linkage study of coeliac disease: failure to replicate previous positive findings. *Ann Hum Genet* 1998;**62**:25-32.

16 **Houlston RS**, Tomlinson IP, Ford D, Seal S, Marossy AM, Ferguson A, Holmes GKT, Hosie KB, Howdle PD, Jewell DP, Godkin A, Kerr GD, Kumar P, Logan RFA, Love AHG, Johnston S, Marsh M, Mitton S, O'Donoghue D, Roberts A, Walker-Smith JA, Stratton MF. Linkage analysis of candidate regions for coeliac disease genes. *Hum Mol Genet* 1997;**6**:1335-9.

17 **Susi M**, Holopainen P, Mustalahti K, Maki M, Partanen J. Candidate gene region 15q26 and genetic susceptibility to coeliac disease in Finnish families. *Scand J Gastroenterol* 2001;**36**:372-4.

18 **Holopainen P**, Mustalahti K, Uimari P, Collin P, Maki M, Partanen J. Candidate gene regions and genetic heterogeneity in gluten sensitivity. *Gut* 2001;**48**:696-701.

19 **Greco L**, Babron MC, Corazza GR, Percopo S, Sica R, Clot F, Fulchignoni-Lataud MC, Zavattari P, Momigliano-Richiardi P, Casari G, Gasparini P, Tosi R, Mantovani V, De Virgiliis S, Iacono G, D'Alfonso A, Selinger-Leneman H, Lemainque A, Serre JL, Clerget-Darpoux F. Existence of a genetic risk factor on chromosome 5q in Italian coeliac disease families. *Ann Hum Genet* 2001;**65**:35-41.

20 **Walker-Smith JA**, Guandalini S, Schmitz J, Schmerling DH, Visakopri JK. Revised criteria for the diagnosis of coeliac disease. *Arch Dis Child* 1990;**65**:909-11.

21 **Olerup O**, Aldener A, Fogdell A. HLA DQ-A1 and DQB1 typing by PCR amplification with sequence specific primers (PCR-SSP) in 2 hours. *Tissue Antigens* 1993;**41**:119-34.

22 **Kruglyak L**, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;**58**:1347-63.

23 **Gudbjartsson DF**, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. *Nat Genet* 2000;**25**:12-13.

24 **Pena AS**, Mann DL, Hague NE, Heck JA, van Leeuwen HA, van Rood JJ, Strober W. Genetic basis of gluten-sensitive enteropathy. *Gastroenterology* 1978;**75**:230-5.

25 **Greenberg DA**, Hodge SE, Rotter JI. Evidence for recessive and against dominant inheritance at the HLA-"linked" locus in coeliac disease. *Am J Hum Genet* 1982;**34**:263-77.

26 **Hernandez JL**, Michalski JP, McCombs CC, McCarthy CF, Stevens FM, Elston RC. Evidence for a dominant gene mechanism underlying coeliac disease in the west of Ireland. *Genet Epidemiol* 1991;**8**:13-27.

ECHO..

Covert operations

Gene therapy holds promise for treating liver diseases. A review in *Gut* by Schmitz *et al* describes significant progress in search of new options based on genetic techniques for treating intractable diseases such as chronic hepatitis B and C infection and primary and secondary liver cancers.

For example, in viral hepatitis gene transfer of IFN α by adenovirus vector prevents the disease in mice, and combined administration of one adenovirus vector containing selected sequences of hepatitis C virus and another containing interleukin 12 genes potentiated a cellular immune response.

In liver cancer drug sensitisation and genetic immunotherapy look promising. Drug sensitisation—transferring a gene for a foreign enzyme to convert a non-toxic prodrug into a anticancer drug—a suicide gene—is best exemplified by the HSV *tk* gene. The encoded enzyme converts the prodrug gancyclovir into a toxic compound which halts DNA synthesis. Added to this is an appreciable bystander effect—when the metabolite diffuses into the surrounding cells, stimulating local inflammation and antitumour immunity, owing to the death of cancer cells. Unwanted effects on healthy tissue are reduced by injecting the vector locally or limiting gene expression with tumour specific promoters.

Genetic immunotherapy offers ways of stimulating the immune system against cancer cells. Injecting a virus vector encoding interleukin 12 into rats with hepatocellular carcinoma in the liver and into mice with metastatic colon cancer achieved complete eradication of tumours and induced antitumour immunity. Introducing this vector through the intrahepatic artery increased antitumour activity in an aggressive hepatocellular carcinoma model induced by a carcinogen.

Therapeutic approaches with HSV *tk* and *interleukin 12* are sufficiently promising for trials in humans to be under way.

▲ *Gut* 2002;**50**:130–135.

Please visit the Journal of Medical Genetics website [www.jmedgenet.com] for link to this full article.